R E P O R T R E S U M E S

ED 017 920 48 AL 001 123

THE CLOZE PROCEDURE--A REVIEW OF THE LITERATURE AND POSSIBLE APPLICATIONS TO THE STUDY OF DEVIANT LANGUAGE FUNCTIONS.

BY- BENNETT, STANLEY AND OTHERS
MICHIGAN UNIV., ANN ARBOR, CTR.FOR RES.LANG.AND BEH

REPORT NUMBER BR-6-1784 PUB DATE 67
CONTRACT OEC-3-6-061784-0508
EDRS PRICE MF-$0.25 HC-$0.68 15P.

DESCRIPTORS- #CLOZE PROCEDURE, #RESEARCH NEEDS, #MENTALLY HANDICAPPED, RESEARCH REVIEWS (PUBLICATIONS), ANALYSIS OF VARIANCE, LANGUAGE ABILITY,

THE "CLOZE PROCEDURE" AS DEFINED BY TAYLOR (1953) IS A TECHNIQUE IN WHICH A MESSAGE IS MUTILATED BY DELETING EVERY N-TH ELEMENT. THE TASK REQUIRES RECONSTRUCTION OF THE MESSAGE IN ITS ORIGINAL FORM. TAYLOR POSTULATED THAT THE "CLOZE UNIT" IS DEPENDENT ON TOTAL CONTEXT AND THE DEGREE OF ASSOCIATION WITH THE IMMEDIATELY SURROUNDING ELEMENTS. A REVIEW OF THE LITERATURE HAS REVEALED A LACK OF EVIDENCE REGARDING THE FUNCTIONING OF MENTALLY RETARDED CHILDREN ON THE CLOZE TASK. THE PRESENT REVIEW OF RESEARCH EXPLORES THE VARIABLES WHICH AFFECT CLOZE PERFORMANCE. THE GOAL OF FUTURE STUDIES IN THIS PROJECT WILL BE TO APPLY CLOZE PROCEDURE IN A RESEARCH PROGRAM DESIGNED TO STUDY THE LANGUAGE FUNCTIONING OF MENTALLY RETARDED CHILDREN. THIS REPORT WAS PUBLISHED IN "STUDIES IN LANGUAGE AND LANGUAGE BEHAVIOR, PROGRESS REPORT IV," 1967, BY THE CENTER FOR RESEARCH ON LANGUAGE AND LANGUAGE BEHAVIOR, UNIVERSITY OF MICHIGAN, 220 EAST HURON STREET, ANN ARBOR, MICHIGAN 48108. (AUTHOR/JD)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

BR-6-1784
PA 48

ED017920

The Cloze Procedure: A Review of the Literature and Possible Applications to the Study of Deviant Language Functions

Stanley Bennett, Melvyn I. Semmel, and Loren S. Barritt
Center for Research on Language and Language Behavior

The "cloze procedure" (CP) was defined by Taylor (1953) as a technique in which a message is mutilated by deleting every nth element. The task requires reconstruction of the message in its original form. Taylor postulated that the "cloze unit" is dependent on total context and the degree of association with the immediately surrounding elements.

A review of the literature has revealed a lack of evidence regarding the functioning of mentally-retarded children on the cloze task (Jordan, 1963; McCarthy, 1964). The present review of research explores the variables which affect cloze performance. Our goal is to apply CP in a research program designed to study the language functioning of mentally retarded children.

The Cloze Procedure as a Measure of Readability

Taylor (1953), whose interest in CP concerned specifically its value as a test of readability, has hypothesized that words deleted from difficult passages are more difficult to replace than words removed from easier passages. He contended that CP measures the similarity between the patterns that the decoder is anticipating and the patterns that the encoder used.

Initially, Taylor compared the amount of agreement between cloze scores and two common measures of readability (Flesch and Dale-Chall). He was also interested in the effect of different patterns of word deletion (e.g., random deletion vs. eliminating every nth word; removing "few" vs. "many" words). Cloze scores yielded rankings of passages on relative difficulty which were similar to those obtained from the Flesch and Dale-Chall formulas. Random deletion patterns and every nth systems resulted in comparable ratings if more than 16 blanks per passage were supplied. No appreciable increase in efficiency of differentiation was obtained by giving partial credit for relevant synonyms substituted for deleted words.

Taylor also tested the power of CP in differentiating passages of prose which differed, by a priori judgment, in reading difficulty. His hypotheses were that the readability of eight previously-rated passages would dependably predict the rank order of the cloze scores; that CP would successfully differentiate the eight passages, whereas the standard (Flesch and Dale-Chall) formulas would not; and that the difference between cloze scores for the various

AL 001 123

passages would be significant. The internal consistency of the cloze scores was also measured by dividing the Ss for each passage into subgroups so that performance of those in the upper third (as judged by total scores) could be compared with the performances of those in the lower third.

The results verified the discriminating power of CP. Ranking of readability by six judges for the eight passages yielded a reliable prediction of the order of difficulty obtained from the cloze scores. These scores came closer than either of the standard formulas in ranking the relative readability levels of the passages. Analysis of variance indicated significant differences in the scores for the eight passages ($p < .01$). Internal consistency coefficients were high enough to indicate that half as many blanks could have been used without changing the ranking of passages. In every case, the cumulative score of the high scorers quickly diverged from that of the low scorers--usually within the first five blanks. Taylor suggested that CP might provide a method for contrasting the reading abilities of different individuals, but that the potential usefulness of the technique is in no way confined to the study of reading.

In another study Taylor (1954) used a 175-word passage from which every fifth word was deleted. Five different cloze versions were constructed by starting with a different word so that all words were deleted in one of the five cloze tests. Each of the cloze versions was given to a group of 50 Ss. Cloze and entropy (degree of uncertainty) scores were determined for each of the overlapping forms. The rank-difference correlation of -.87 between cloze and entropy scores was interpreted to mean that blanks having high uncertainty yield significantly lower cloze scores.

The Cloze Procedure as a Measure of Reading Comprehension

Taylor (1956) reported a study designed to measure how well Ss understood the meaning of a passage. It was hypothesized that, as understanding increased, the probability of providing the correct word increases and, consequently, the probability of restoring the passage to its original form is enhanced. Taylor also contended that cloze scores would predict intelligence, existing knowledge, and success in learning and remembering. Air Force trainees were used as Ss. Comprehension tests for an Air Force technical article were administered before and after completion of the article. Cloze tests were constructed by using a

sample representing 20 per cent of the article. Ten words were "mechanically" deleted from each nine lines. The Ss were first given the cloze test and then the "before" comprehension test. A week later they read the technical article which was immediately followed by administration of the "after" comprehension test, which in turn was followed by another form of the cloze test. Correlations between cloze scores, comprehension measures and intelligence test scores were computed. Both "before" and "after" cloze scores correlated highly with the comprehension scores. The intelligence measure and the "before" cloze test scores correlated .73. The "after" cloze and intelligence test scores correlated .74. The cloze scores predicted "learning" after study as adequately as comprehension test scores.

Three deletion patterns were used in another study reported by Taylor (1957): (1) deletion of any word; (2) deletion of "hard" words (nouns, verbs, adverbs); and (3) deletion of "easy" words (verb auxiliaries, conjunctions, pronouns, and articles). The "ease" in filling in these words was determined by the author from previous studies. The results indicated that in general the "any" cloze form yielded more stable, reliable and discriminating results than did the "easy" or "hard" forms. The "any" deletion pattern yielded correlations with the comprehension test of between .70 and .88; in all but one case these coefficients were larger than those for the "hard" or "easy" deletion patterns. The findings from this experiment confirm cloze readability scores as valid measures of the comprehension of English prose.

The Cloze Procedure as a Measure of Reading Proficiency

The initial studies by Taylor on the relationship between cloze performance and readability were quickly followed by a series of studies relating cloze scores to individual reading proficiency. Notable among these were studies reported by Jenkinson (1957), Rankin (1958), Fletcher (1959), Bormuth (1962), and , more recently, Greene (1964).

Jenkinson (1957) studied CP responses of high school students to three deletion conditions: (1) key words (inferred from the surrounding context); (2) words from three form classes; and (3) 90 per cent of the high-frequency words. Jenkinson found correlations between total cloze scores and the Cooperative Reading Test of .78 and .73. Significant differences in cloze scores across grade levels were paralleled by reading test scores.

Rankin (1958) employed CP to assess gains in reading skill after a reading improvement course. Five cloze tests were developed. In one form, every fifth word was deleted. The other four forms deleted nouns and verbs alternately. Cloze scores were found to be significantly related to intelligence. Ss who received higher intelligence scores had higher pre- and post-test reading scores. However, the amount of gain was not related to IQ scores. Reading rate was not related significantly to post-reading cloze scores for the passages in which nouns and verbs had been deleted.

Fletcher (1959) related cloze scores to reading comprehension, reading rate and general verbal ability. Every fifth word was deleted, irrespective of form class. He found correlations of .59 and .57 between the cloze scores and two tests of reading rate. On the comprehension test, the correlation with the cloze score was .55. A test of general verbal ability (ACE) correlated .72 with the CP scores. Fletcher also found that content words were more difficult to supply than function words. High-frequency words were the least difficult to supply and had little tendency to occur among the items that discriminated best.

Fletcher observed that the CP responses were similar to those frequently observed in controlled association studies, (i.e., a small number of different responses were given to each item with one word given a large proportion of the time). Low scorers left more items blank and tended to give uncommon responses than did high scorers.

Three aspects of CP were investigated by Bormuth (1962): (1) validity as a measure of comprehension ability; (2) efficiency in discriminating between difficulties of tests and abilities of individuals; and (3) validity as a measure of the amount of comprehension with which passages are read.

Comprehension tests were developed for the nine passages used. Cloze tests for each passage were constructed by deleting every fifth word. Only exact replacement of the deleted word was scored as correct. Children from grades 4, 5, 6 were the Ss.

The results support the hypothesis that cloze tests are related to measures of comprehension. All correlations between comprehension and cloze scores were significant. Item reliability for the cloze test was sufficient to insure efficiency in discriminating both tests and individuals.

Greene (1964) conducted a study in which both the standard cloze technique (deleting every nth word) and a modified method were used. In the modified procedure, the only words eligible for deletion were nouns, verbs, adjectives, and adverbs. In this latter condition, each deleted word was evaluated for its value as a test item by examining the redundancy remaining in the passage and selecting those items which it was felt could be replaced by a superior reader.

Greene was specifically interested in the relative efficiency of the two deletion procedures in predicting scores on standardized reading achievement tests. No significant differences were found between the two procedures as measured by their correlations with reading tests: both cloze techniques correlated highly with the reading tests. The modified cloze test was significantly more reliable than the standard CP. Contrary to his prediction, Greene found reading speed to be significantly related to cloze test performance. Two vocabulary tests accounted for approximately one-third of the total cloze-test variance. Thus, after taking into consideration the variables of vocabulary and rate of reading, a large proportion of the cloze test variance remained unaccounted for.

It has already been noted that Taylor (1956), Jenkinson (1957) and Rankin (1958) found that intelligence was an important variable in cloze performance. Ruddell (1963) constructed three cloze tests with sentences of high frequency and three with sentences of low frequency in oral speech. No significant differences were found in the performance of fourth-graders on the high and low frequency tests. However, the following factors did contribute to performance on both kinds of cloze tests: (1) father's occupation; (2) mother's educational background; (3) intelligence; (4) mental age; and (5) chronological age.

Weaver (1961) presented cloze passages aurally or by having Ss read and complete the passages. Cloze tests were constructed with differing criteria for deletion: (1) every nth word; and (2) deletion of content words (nouns, verbs). He found that the "any-word" pattern of deletion was easier to solve when read by the S, but that the content-word pattern of deletion yielded similar results in both the listening and reading conditions. Weaver suggests that a structural deletion pattern (every nth word) is more difficult when S

listens (instead of reading) because the information sequence and mode of presentation are under the control of the speaker rather than the S.

Effect of Amount and Type of Context on Cloze Performance

Next to be considered is a series of studies dealing with the effect of the amount and type of context on cloze proficiency. Although it did not involve deletion of words, a study by Burton and Licklider (1955) is relevant to the discussion. These authors studied the extent to which estimates of the redundancy of English texts are dependent upon the number of preceding letters known to the S. They found that, although relative redundancy increased as knowledge of the foregoing text approached 32 letters, providing more than 32 letters did not result in a noticeable increase in performance.

Miller and Friedman (1957) eliminated or substituted letters in printed English texts, added extra letters, etc. They found that "average" Ss were unable to correct such passages if more than 10 per cent of the letters were altered. Random substitution of letters was clearly the most difficult condition. Superior Ss with unlimited time were able to correct passages with as much as 50 per cent of the text "mutilated".

Using complete prose paragraphs, MacGinitie (1961) studied the effects of different patterns of omission on the difficulty of restoring periodically omitted words. He found that words were equally restorable when every 24th, 12th or 6th word was omitted but, when every third word was omitted, restoration was considerably more difficult. Words omitted in groups of four were more difficult to replace than words omitted in pairs, and pairs were more difficult than single words. MacGinitie concluded that context more than five words distant from the omitted word has relatively little effect upon cloze test performance.

The results of the MacGinitie study appear to be related to the Burton and Licklider study in that both indicated a basic limitation in the length of units which can be restored from context. The redundancy of our language is primarily effective for small segments of a message. That is, if omitted segments are small they can be restored with great accuracy. If groups of words or letters are omitted, the loss is usually irretrievable from context, even though such groups may be few and widely separated.

Shepard (1963) required Ss to produce a list of words which could be used to fill a blank representing a word deleted from a passage of text. The amount of context that originally surrounded the deleted word was varied. The hypothesis tested was that as the amount of context was increased, the number of alternative words that could fit in the blank decreased. The amount of context used was 1,2,4,6,10, or 40 words (as they occurred in original passages in the New York Times). Ss were given five minutes to list as many words as possible that might be placed in the missing blank. Shepard found that the average rate of responding fell from about 20 words per minute, when the smallest amount of context was provided, to two words per minute when the maximum amount of context was given. Shepard stated that, since rate of producing words may be a function of uncertainty associated with a category, uncertainty might be inferred from rate.

Effect of Form Class on Cloze Proficiency

Another series of studies has been devoted to testing the effects of form class on the relative ease in filling in cloze passages. Both Jenkinson (1957) and Fletcher (1959) found that content words were more difficult to replace than function words.

Fillenbaum, Jones, and Rapoport (1963) investigated the effect of rate of deletion upon predictability of form class replacement in a cloze task. A sample of continuous speech elicited by TAT cards, was used with deletion of every second, third, fourth, fifth, or sixth word. Each deleted word and each word supplied by the S was classified into one of eleven grammatical categories.

The results showed that, with every second word deleted, the form class of the missing word was correctly identified in nearly two-thirds of the cases. The largest increase in predictability occurred between deletion rates of every two and every three words. When the data were scored for exact agreement between the words of the original passage and responses of the Ss (verbatim responses), adjectives were found to have been most difficult to replace. In form class identification, Ss had most difficulty with adverb and quantifier slots. Nouns were easiest to replace, followed by prepositions, then pronouns, auxiliary verbs, adjectives, and articles. In general, "syntactic words" (prepositions, conjunctions, articles, auxiliary verbs) were more easily replaced than "semantic words" (nouns, verbs, and adjectives).

One hundred college students were tested by Coleman and Blumfeld (1963) using nominalized sentences (e.g., "our goal is the achievement of the highest good for society, to be attained through a correct definition of names.") and grammatical transformations of these sentences with active verbs (e.g., "Our goal is that we achieve the highest good for society and this is attained when we define names correctly."). Every fifth word was eliminated and replaced by a blank 10 spaces long. On the average, 9.63 words per S were supplied correctly in the nominalized sentences and 10.80 in the sentences using active verbs. The mean number of blanks correctly filled in for content words was 1.44 for nominalized sentences, in contrast to 2.22 for active verb sentences; a significant difference ($p < .01$). Proficiency in filling in function words was not significantly different for the two types of sentences.

An analysis of correct responses by form-class revealed that nouns, verbs, adjectives and adverbs were equally difficult to replace, whereas pronouns and function words were more easily supplied. Whatever the form class, Ss were able to fill in more blanks in sentences with active verbs than in nominalized sentences.

The percentages of correct responses by form class are lower than those found by Aborn, Rubinstein, and Sterling (1959), who were interested in the constraint upon words attributable to the amount, distribution, and structure of context in cloze sentences. Sentences of 6, 11, and 25 words each were drawn from popular magazines. Each word of every sample was classified according to form class: i.e., noun, verb, adjective, adverb, pronoun, or function word. Function words included articles, conjunctions, prepositions, auxiliary verbs, interjections, and quantity words. Words were deleted in one of four positions: (1) initial position; (2) early medial (the third word in six-word sentences, the fourth word in 11-word sentences, and the eighth word in 25-word sentences); (3) late medial (fourth word in six-word sentences, eighth word in 11-word sentences, and the 17th word in 25-word sentences); and (4) final position. Sentences were chosen so that an equal number began with words of each of the six different form classes.

Function words and pronouns were found to be considerably more predictable than words of the other form classes. For the most part, the percentage of words correctly given was a function of the size of the class, i.e., the smaller the frequency of words in that class, the greater the probability of correct prediction. The descending hierarchy of classes according to size would appear from

these data and those of Fries (1952) to be as follows: nouns, verbs, adjectives adverbs, function words, and pronouns. The tendency to draw responses from a class other than that of the omitted word was most pronounced in the case of the adjective and adverb.

In analyzing for the position effect, it was found that words omitted in the medial positions were more predictable than words omitted in the initial or final positions. Thus, as reported by Miller and Friedman (1957), bilateral context exerts greater constraint than unilateral context (regardless of whether the context precedes or follows the missing word). The results indicate that predictability of the missing word increases with increasing sentence length, but that this reaches an asymptote, with no noticeable increase in predictability for the 25-word sentences over the 11-word sentences. It thus appears that constraint reaches its maximum somewhere between sentence lengths of five and ten words.

Contribution of Syntactical and Semantic Components to Cloze Performance

In 1961 MacGinitie stated that: "Although various sources of constraint (e.g., topic, syntax, semantic reasonableness) can be distinguished conceptually, it has long been recognized that it is difficult in completion experiments to separate their effects" (p. 128). A study by Salzinger, Portnoy, and Feldman (1962) attempted to investigate the semantic and syntactic factors in verbal behavior using CP.

Fifty-word passages developed by Miller and Selfridge (1950) at each of eight orders of statistical approximation to English (zero, first...seventh, text) were used. Either every fifth or every seventh word was deleted by Salzinger et al. The results indicated that as the order of approximation approached normal English, more deleted words were correctly supplied. Further, these authors found, as did MacGinitie (1961) and Aborn, Rubinstein, and Sterling, (1959), that Ss do not or cannot make use of more than five words on either side of each blank. By classifying both the incorrect and correct cloze responses, it was discovered that, with increasing order of approximation to English, Ss gave increasingly more words of the same form class as the deleted word; however, the effect did not increase appreciably beyond Order 3. Salzinger et al. suggest that improvement in memory for sentences must be attributed: primarily to increased syntactical structure between Orders 1 and 2, about equally to syntax and meaning between Orders 2 and 3, and primarily to meaning beyond Order 3.

Miscellaneous Variables in Cloze Performance

Manis and Dawes (1961) studied the variable of attitudes toward the "mutilated" and found that passage responses to blanks formed by deletion of content words did not reflect differences attributable to previously-expressed agreement or disagreement with the passage, while responses to function words were affected by prior opinion.

Fillenbaum and Jones (1962) report a study which attempts to measure, via CP, the extent of deviation of aphasic speech from normal speech. Cloze tests were constructed by deleting every fifth word from speech transcripts of normal and aphasic Ss. College students then were given these cloze tests, and the number of correctly-replaced words in the normal transcript was compared to that in the aphasic transcript.

Cloze scores were significantly higher for passages from "normal" speech than from aphasic speech. Even if a rater's completion of the cloze blank was of the correct form class, the probability that his verbatim prediction was correct was consistently less for passages generated by aphasics than for passages taken from normals. Two cloze scores, form class predictability and verbatim predictability, clearly discriminated between texts; these scores appear to have contributed independently to such discrimination.

Fillenbaum and Jones suggest that the conditional proportion of correct verbatim completions given correct form-class substitution might serve as an especially powerful index of the amount of semantic difficulty of an aphasic speaker. Low predictability of form class would suggest syntactic problems. The authors stated that it might be possible to examine the predictability of different form classes, thereby specifying more precisely the locus of change or damage in speech.

Summary and Discussion

Since the early readability studies by Taylor (1953, 1954), much progress has been made in delineating the variables contributing to cloze performance. Taylor (1956, 1957) found that cloze scores correlated significantly with tests of reading comprehension and general intelligence. Jenkinson (1957), Rankin (1958), Fletcher (1959), Bormuth (1962) and Greene (1964) all found significant correlations between individual reading ability and cloze performance. The effect of intelligence on cloze proficiency was also confirmed.

Ruddell (1963) found socio-economic class, mental age and chronological age to be significant factors in performance on cloze passages.

Another series of studies was concerned with the effects of amount of context, form class of words deleted and distribution of the deleted items. Burton and Licklider (1955), Shepard (1963), and Fillenbaum, Jones and Rapoport (1963), found that the amount of available context affected ability to fill in cloze blanks. Jenkinson (1957), Fletcher (1959), Aborn, Rubinstein, and Sterling (1959), and Coleman and Blumfeld (1963), and Fillenbaum, Jones, and Rapoport (1963), discovered that deleted function words were easier to supply than content words. This has been explained on the basis of the number of words in the form class, i.e., there are comparatively fewer function words than content words, resulting in fewer alternatives and therefore greater predictability. Aborn, Rubinstein, and Sterling (1959) also found that the distribution of the deleted word (deletion in initial, medial or final positions of the sentence) affected the ability to correctly supply the missing word. It was shown that blanks in the initial or final positions were more difficult to fill than those in medial positions. Aborn, et al. concluded, as had Miller and Friedman (1957), that bilateral context exerts greater constraint than does unilateral context.

Salzinger, Portnoy, and Feldman (1962) attempted to differentiate the effects of syntax and semantics in the recall of sentences of different orders of approximation to English. Using the Miller and Selfridge 50-word passages, Cloze tests were constructed to evaluate the nature of responding at each of the eight orders of approximation to English. It was found that with increasing order of approximation, Ss moved from a reliance on syntactical cues to a semantic level of functioning. Manis and Dawes (1961) determined that the effect of agreement or disagreement with the passage read affects mostly the function words.

The Fillenbaum and Jones (1962) study presents evidence that aphasic speech may be objectively evaluated by means of the cloze technique. This study with aphasics highlights the potential of the cloze technique in dealing with language disorders. It may be possible to apply the cloze procedure to the study of verbal development in retarded children. Perhaps it would be possible to determine the extent to which retardates have learned the "equivalence" of words from the same form class. In fact, it might be possible

to make a prediction as to the relative proficiency on the cloze test based upon the amount of paradigmatic responding (Semmel, Barritt, Bennett, and Perfetti, 1966) on a free-association task. It can be hypothesized that Ss giving paradigmatic responses will have in their repertoires the largest store of available "equivalent" words substitutable for the missing word. This would result in higher cloze scores when scoring was done on the basis of whether the substituted word was of the same form class as the deleted words. This procedure promises to yield data concerning the amount of information retardates obtain from the grammatical structure of connected discourse.

Another factor of significance may be the extent to which S is capable of paradigmatic responding. The pattern of associations between individual words may help to explain cloze responses. Some support for the organizing effect of associational bonds and the consequent effect on cloze performance is obtained in the Fletcher (1959) study, where it was found that high scorers showed more commonality of responding than low scorers. Another way of scoring the cloze data would be for exact (verbatim) agreement between the word deleted and the word supplied by the S. Fillenbaum and Jones (1962) suggest that this may give some indication as to the level of semantic development.

One variable that has not been studied is the effect of different levels of complexity upon cloze responses. It would seem appropriate to begin a study of this kind at the phrase level, using various types of phrases (e.g., noun, verb, and prepositional phrases), noting the relative ease with which missing words can be supplied. Then simple sentences could be formulated, containing different orderings of the constituent form classes (e.g., "Go to the store," "The home is white,", etc.). Thus, not only would it be possible to judge the relative difficulty Ss have in supplying words of certain form classes but also the impact of particular form-class orders. Next, cloze performance on phrases and simple, disconnected sentences could be contrasted with performance on continuous prose. It might be expected that retarded children and children with speech disorders would have progressively more difficulty with the cloze task in moving from the phrase level to continuous prose, partially because of increasing constraint due to context.

We might also speculate that because of greater familiarity with nouns and verbs, retarded Ss would be most adept at filling in cloze blanks requiring these form classes. It seems likely from previous studies that cloze performance will be related to both chronological age and mental age, irrespective of the level of language sampled.

The efficiency of the cloze procedure in measuring comprehension of a passage of prose should be evaluated by correlating cloze scores with different types of comprehension tests (e.g., essay, short answer, multiple-choice). It may be that the high correlations found between comprehension tests and cloze tests are an artifact of the type of test given. It would seem on an a priori basis that multiple-choice and short-answer tests would correlate more highly with the cloze task than would an essay test.

Taylor (1953) said this about the cloze procedure: "...a cloze score appears to be a measure of the aggregate influences of all factors which interact to affect the degree of correspondence between the language patterns of transmitter and receiver. As such, its potential usefulness is by no means confined either to readability or the reading abilities of individuals" (p. 432).

Future research with the cloze procedure should attempt to further differentiate the contributing "aggregate" influences mentioned by Taylor. The use of the cloze technique with new populations may not only solve important theoretical issues, but may also prove to be a valuable index of verbal functioning with important predictive characteristics.

References

Aborn, M., Rubinstein, H., & Sterling, T. D. Sources of contextual constraint upon words in sentences. *J. exp. Psychol.*, 1959, *57*, 171-180.

Bormuth, J. R. Cloze tests as measures of readability and comprehension ability. Unpublished doctoral dissertation, Indiana Univ., 1962.

Burton, N. G., & Licklider, J. C. R. Long-range constraints in the statistical structure of printed English. *Amer. J. Psychol.*, 1955, *68*, 650-653.

Carroll, J. B., Carton, A. S., & Wilds, C. P. An investigation of "cloze" items in the measurement of achievement in foreign languages. Cambridge, Mass: Laboratory for Research in Instruction, Graduate School of Education, Harvard Univ., 1959.

Coleman, E. B., & Blumfeld, J. P. Cloze scores of nominalizations and their grammatical transformations using active verbs. *Psychol. Rep.*, 1963, *13*, 651-654.

Dale, E., & Chall, Jeanne. A formula for predicting readability. *Educ. Res. Bull.*, 1948, *27*, 11-20 and 37-54.

Ervin, Susan. Changes with age on the verbal determinants of word association. *Amer. J. Psychol.*, 1961, *74*, 361-372.

Fillenbaum, S., & Jones, L. V. An application of "cloze" technique to the study of aphasic speech. *J. abnorm. soc. Psychol.*, 1962, *65*, 183-189.

Fillenbaum S., Jones, L. V., & Rapoport, A. The predictability of words and their grammatical classes as a function of rate of deletion from a speech transcript. *J. verb. Learn. verb. Behav.*, 1963, *2*, 186-194.

Flesch, R. *How to Test Readability*. New York: Harpers, 1951.

Fletcher, J. E. A study of the relationships between ability to use context as an aid in reading and other verbal abilities. Unpublished doctoral dissertation, Univer. of Washington, 1959.

Greene, F. P. A modified cloze procedure for assessing adult reading comprehension. Unpublished doctoral dissertation, Univer. of Michigan, 1964.

Jenkinson, M. D. Selected processes and difficulties of reading comprehension. Unpublished doctoral dissertation, Univer. of Chicago, 1957.

Luke, Molly. Form class and cloze procedure. Unpublished study, Language Development Program, Center for Human Growth and Development, Univer. of Michigan, 1964.

MacGinitie, W. H. Contextual constraint in English Prose paragraphs. *J. Psychol.*, 1961, *1951*, 121-130.

Manis, M., & Dawes, R. M. Cloze scores as a function of attitude. *Psychol. Rep.*, 1961, *9*, 79-84.

Miller, G. A., & Friedman, E. A. The reconstruction of mutilated English texts. *Information and Control*. 1957, *1*, 38-55.

Miller, G..& Selfridge, Jennifer. Verbal context and the recall of meaningful material. *Amer. J. Psychol.*, 1950, *63*, 176-185.

Rankin, E. F. An evalution of the cloze procedure as a technique for measuring reading comprehension. Unpublished doctoral dissertation, Univer. of Michigan, 1958.

Ruddell, R. B. An investigation of the effect of the similarity of oral and written patterns of language structure on reading comprehension. Unpublished doctoral dissertation, Indiana Univer., 1963.

Salzinger, K., Portnoy, Stephanie, & Feldman, R. S. The effect of order of approximation to the statistical structure of English on the emission of verbal responses. *J. exp. Psychol.*, 1962, *64*, 52-57.

Shepard, R. N. Production of constrained associates and the informational uncertainty of the constraint. *Amer. J. Psychol.*, 1963, *76*, 218-228.

Taylor, W. L. "Cloze procedure": a new tool for measuring readability. *Journalism Quart.*, 1953, *30*, 415-433.

Taylor, W. L. Application of "cloze" and entropy measures to the study of contextual constraints in samples of continuous prose. Unpublished doctoral dissertation, Univer. of Illinois, 1954.

Taylor, W. L. Recent development in the use of cloze procedure. *Journalism Quart.*, 1956, *33*, 42-48.

Taylor, W. L. "Cloze" readability scores as indices of individual differences in comprehension and aptitude. *J. appl. Psychol.*, 1957, *41*, 19-26.

Weaver, W. W. An examination of some differences in oral and written language using the cloze procedure. Unpublished doctoral dissertation, Univer. of Georgia, 1961.